

Petr KONEČNÝ¹

SIMULACE KORELOVANÝCH NEPARAMETRICKÝCH ROZDĚLENÍ
V RÁMCI METODY SBRA

Abstrakt

Příspěvek se zabývá ověřením procesu generování korelovaných neparametrických rozdělení, histogramů, v rámci metody SBRA (viz. [4], [3]). Proces generování korelovaných histogramů využívá transformace korelovaných gaussových rozdělení na rozdělení obecná dle [7], [8]. Aplikovatelnosti procesu je ověřena na příkladu naměřených vzorků korelované meze kluzu f_y a meze pevnosti f_u z [9].

1. ÚVOD

Stávající přístup v rámci pravděpodobnostního přístupu metodou SBRA uvažuje většinu proměnných, až na výjimky tzv. „existenční“ závislosti (zatížení mostovým jeřábem, dvou-komponentní větrná ružice), jako vzájemně nezávislé. Publikace Tereco [3] sice uvádí jistý přístup pro tvorbu korelovaných obecných rozdělení, který je ovšem z pohledu autora nedostatečný. Generování korelovaného normálního (gaussova) rozdělení je poměrně známý proces, kterým se pro účely aplikace v Anthillu zabýval rovněž Menčík [6].

Generování obecných rozdělení je ovšem náročnější oříšek. Tento problém se pro menší počet simulací podařilo vyřešit za pomoci tzv. simulovaného žihání brněnský kolektiv, kolem prof. Nováka [2] a [10]. Tento přístup je ovšem pro větší počet simulací nutných pro přímou metodu Monte Carlo nevhodný, neboť je velmi výpočetně náročný.

Nadějným přístupem se jeví metoda prezentována Phoonem [7] a [8]. Tento přístup se opírá o generování obecných rozdělení na podkladě korelačních koeficientů naměřeného souboru, distribučních funkcí jednotlivých naměřených veličin, korelovaného normálního rozdělení a vhodné transformace normálního rozdělení na rozdělení obecné [13].

2. PŘÍSTUP K ŘEŠENÍ

Model pro generování korelovaných rozdělení je aplikován na základě práce Phoon et al (viz. např. [7], [8]) a obohacen o přímou aplikaci na diskretní histogram.

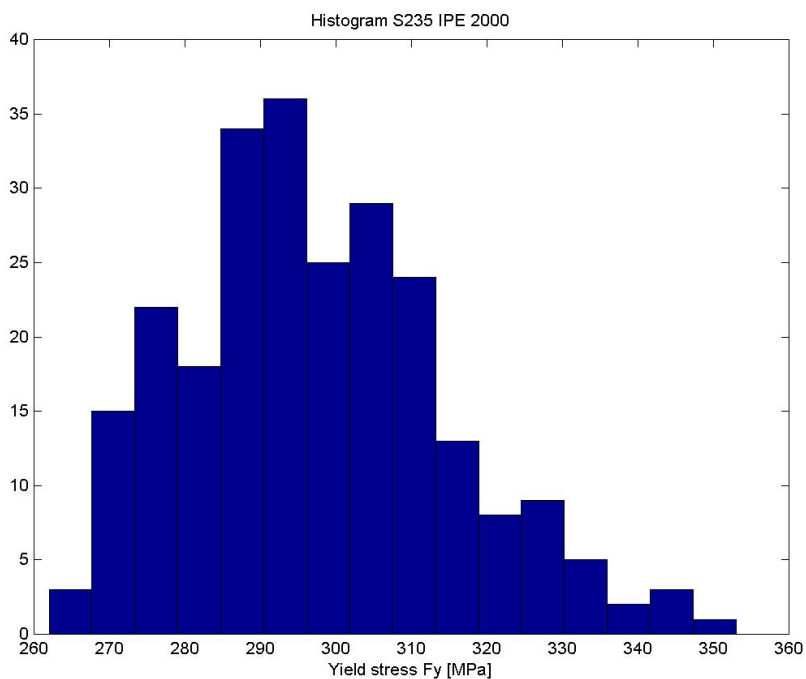
3. NUMERICKÝ MODEL

Zjistíte korelaci mezi daty meze kluzu a meze pevnosti získaných z měření dle [9] a nasimulujte 25 tis. příslušně korelovaných realizací.

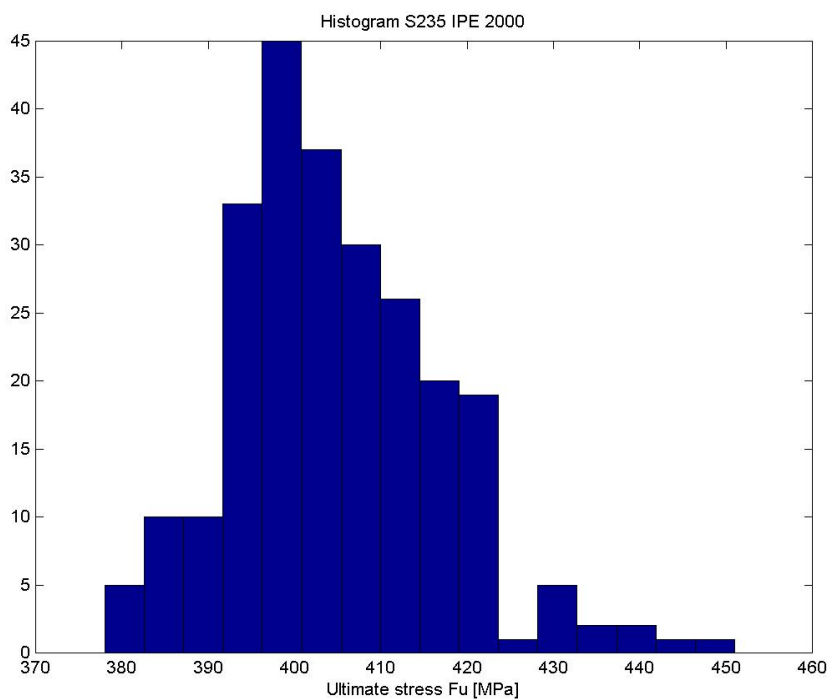
3.1. Načtení dat

Na třech následujících obrázcích jsou zobrazeny histogramy meze kluzu f_y , histogramy meze pevnosti f_u , a vztah mezi veličinami f_y a f_u . Tyto data odpovídají 247 zkoušeným vzorkům dle [9]. Dle seskupení „mraveniště“ na Obr. 1 se dá očekávat, že tato data budou korelována.

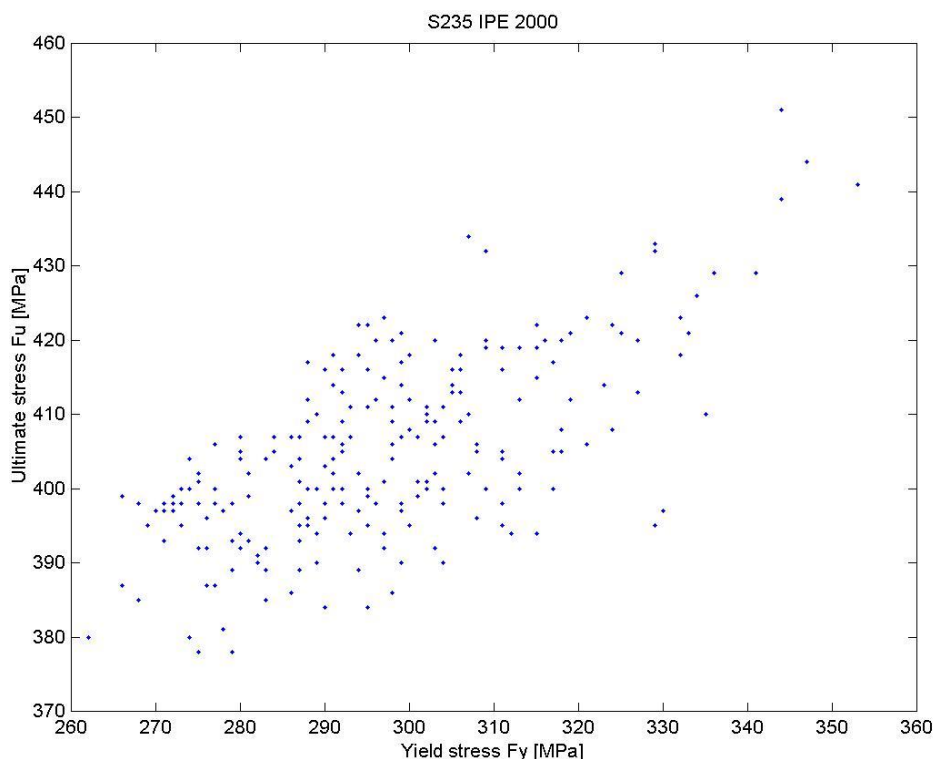
¹ Petr Konečný, Ing., VŠB – TU Ostrava, Fakulta stavební, Katedra stavební mechaniky, L. Poděště 1875, 708 33 Ostrava-Poruba, petr.konecny@vsb.cz



Obr. 1. Histogram meze kluzu f_y [MPa], data dle [9].



Obr. 2. Histogram meze pevnosti f_u [MPa], data dle [9].



Obr. 3. Korelace mezi mezí kluzu f_y a mezí pevnosti f_u [MPa], data dle [9].

3.2. Pearsonův korelační součinitel

Vzájemný vztah f_y a f_u je vyšetřen s využitím Personova korelačního součinitele, zde označeného r_p .

$$r_{p,ij} = \frac{\text{cov}(x_i, x_j)}{s_i s_j} = \frac{1}{s_i s_j} \frac{1}{N-1} \sum_{K=1}^N (x_{i,K} - m_i)(x_{j,K} - m_j) \quad (1)$$

což lze dle [11] vyjádřit jako:

$$r_{p,ij} = \frac{\text{cov}(x_i, x_j)}{\text{var}(x_i) \text{var}(x_j)} = \frac{\sum x_i x_j - N m_i m_j}{\sum x_i^2 - N m_i^2 + \sum x_j^2 - N m_j^2} \quad (2)$$

kde jsou

i mez kluzu f_y ,

j mez pevnosti f_u ,

x_i, x_j korelované vektory,

s_i, s_j směrodatné odchylky příslušných vektorů,

m_i, m_j aritmetické průměry příslušných vektorů,

N je počet prvků ve vektorech i a j .

Následná korelační matice potvrzuje domněnku o korelaci neboť korelace mezi jednotlivými vektory je 0.643:

$$R_p = \begin{bmatrix} 1.0 & 0.643 \\ 0.643 & 1.0 \end{bmatrix} \quad (3)$$

3.3. Transformace korelační matice

Vzhledem k odlišnosti normálního rozdělení a rozdělení obecného charakterizovaného např. histogramem je doporučeno v [7] opravit korelační matici R_p (1) následujícím vztahem za účelem popisu korelace jednotlivých tříd histogramu (fraktilová korelace).

$$r_{ij} = 2 \sin\left(\frac{\pi}{6} r_{p,ij}\right) \quad (4)$$

Fraktilová korelace je tedy popsána následovně:

$$R = \begin{bmatrix} 1.0 & 0.660 \\ 0.660 & 1.0 \end{bmatrix} \quad (5)$$

Získaná korelační matice R musí splnit podmínku pozitivní definitnosti, aby bylo možno realizovat vektory korelovaného normálního rozdělení. Výstupy následujícího testu hledajícího tzv. „eigenvalues“ musí být větší rovny nule.

$$\text{eig}(R) = [0.340 \quad 1.660] \geq 0 \quad (6)$$

Podmínka je splněna a korelační matice R je pozitivně definitní

3.4. Generování korelovaných normálních rozdělení

Korelovaná normální rozdělení je možno dle [13] generovat dvěma způsoby, a to za pomoci vlastních čísel, či dále uvedenou Choleskiho faktorizací.

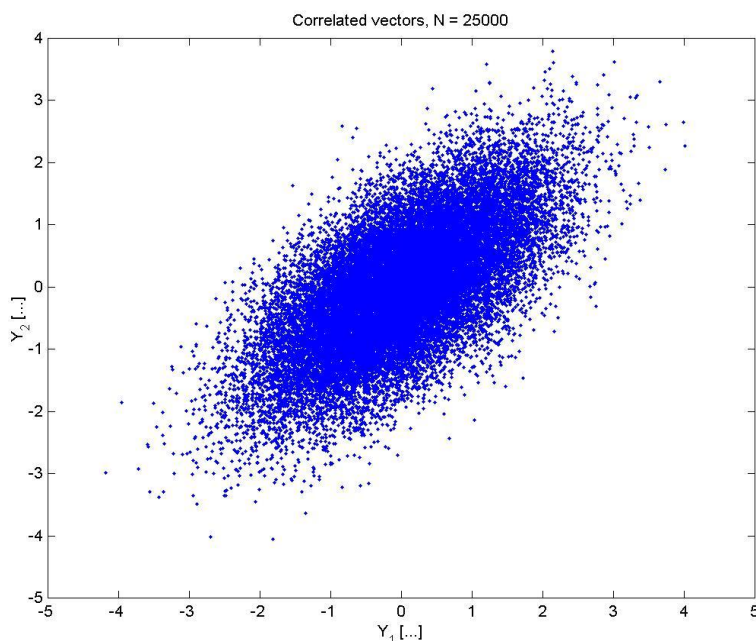
$$Y_{[k,n]} = W_{[k,n]} \times Q_{[k,k]} \quad (7)$$

kde

Y je matice náhodných korelovaných realizací normálního rozdělení,
 W je matice náhodných nekorelovaných realizací normálního rozdělení,
 k počet vektorů (korelovaných proměnných),
 n počet požadovaných realizací (počet simulací) a matice Q je Choleskiho faktorizací korelační matice R splňující podmínku $Q^T \times Q = R$:

$$Q_{[k,k]} = \text{chol}(R) = \begin{bmatrix} 1.000 & 0.660 \\ 0.000 & 0.751 \end{bmatrix} \quad (8)$$

Generování korelovaných normálních rozdělení v programu Anthill [1] je popsáno v Příloze. Na Obr. 4 je zobrazeno 25 tis. realizací korelovaného normálního rozdělení.



Obr. 4. Korelovaná normální rozdělení Y

3.5. Transformace normálního rozdělení na rozdělení rovnoměrné

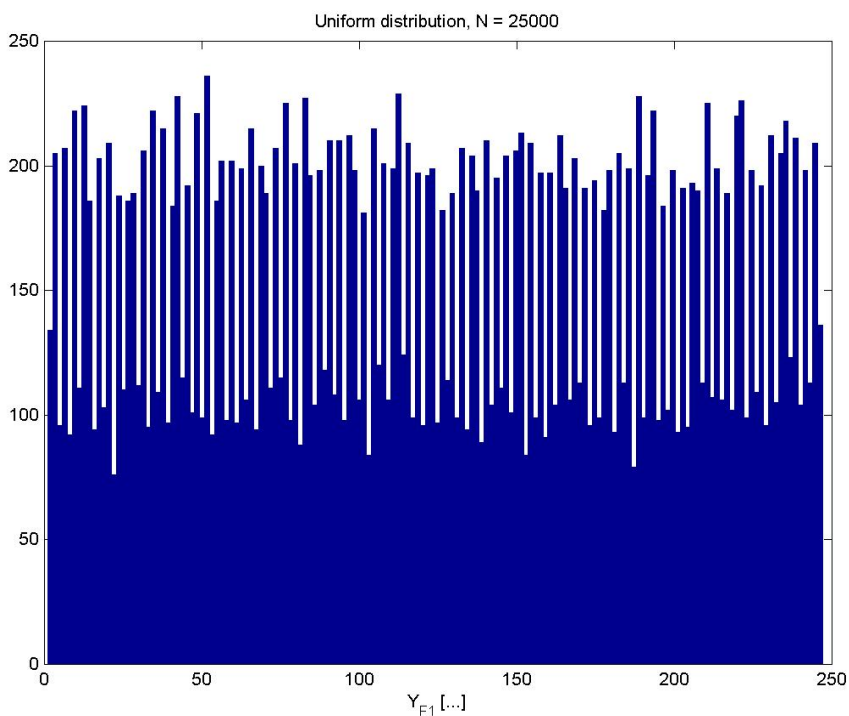
Dva získané vektory normálních rozdělení jež mají vhodnou korelaci je potřeba transformovat na rozdělení rovnoměrné (uniformní) Y_F s využitím distribuční funkce normálního rozdělení (viz. např. [12]):

$$Y_F = \Phi_i = 0.5 + \frac{1}{2} \operatorname{erf}\left(\frac{Y_i}{\sqrt{2}}\right) \quad (9)$$

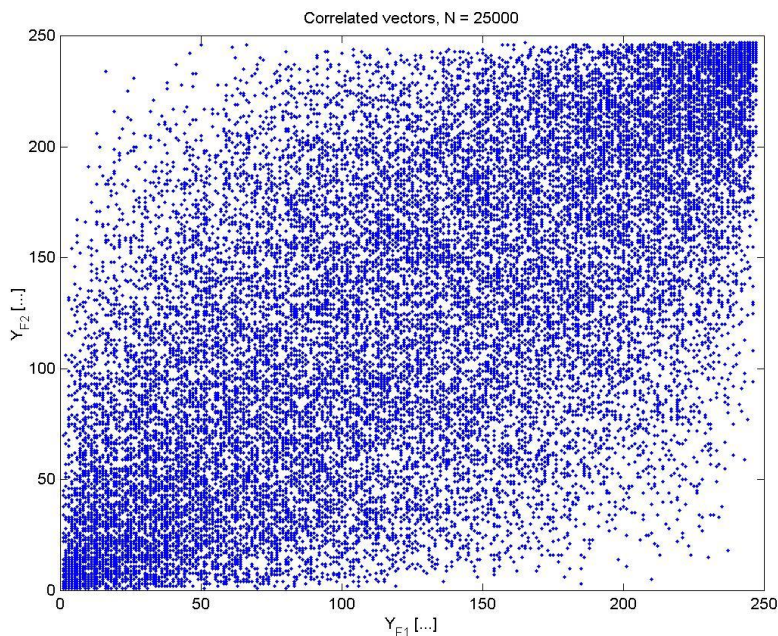
kde erf je tak zvaná „error function“ již lze vypočítat numericky, či rozvinout pomocí vhodného polynomu (pro obor $Y_F < -1.5; 1.5 >$ např.

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \left(x - \frac{x^3}{3} + \frac{x^5}{10} - \frac{x^7}{42} + \frac{x^9}{216} - \frac{x^{11}}{1320} + \frac{x^{13}}{9360} - \frac{x^{15}}{75600} + \frac{x^{17}}{6854440} - \frac{x^{19}}{6894720} + \frac{x^{21}}{76204800} \right) \quad (10)$$

S využitím distribuční funkce normálního rozdělení je získáno rozdělení rovnoměrné, které je vhodné pro generování obecného rozdělení, např. histogramu. Toto rozdělení Y_{F1} je zobrazeno na Obr. 5 (rozdělení pro 2. vektor realizací Y_{F2} vypadá obdobně). Korelace dvou rovnoměrných rozdělení je zobrazena na Obr. 6.



Obr. 5. Upravené rovnoměrné rozdělení získané z normálního transformací dle (9)
 Poznámka: Schodovitost rozdělení je pravděpodobně dána nevhodně zvoleným počtem tříd pro grafické zobrazení rozdělení a nemá vliv na kvalitu procesu, jak naznačuje i další obrázek.



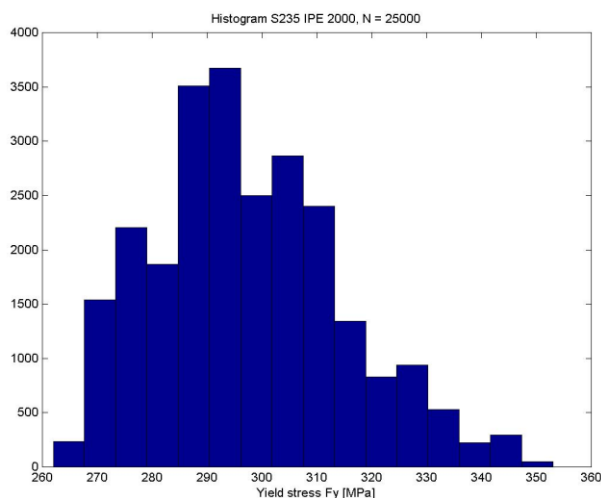
Obr. 6. Korelovaná rovnoměrná rozdělení Y_F

3.6. Transformace rovnoměrného rozdělení na rozdělení obecné

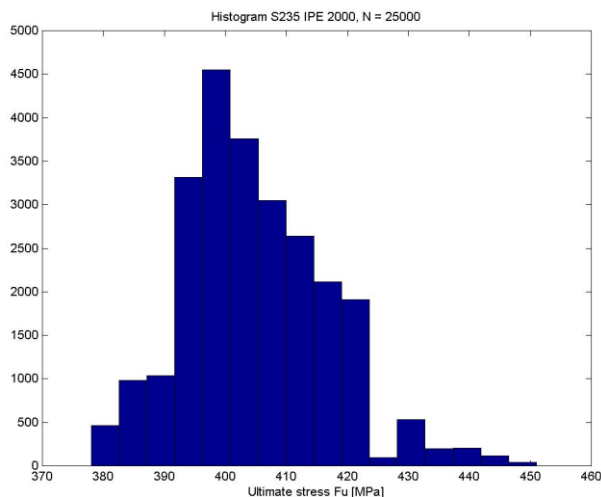
Příslušné neparametrické rozdělení je následně získáno za pomoci inverzní transformace distribuční funkce hledaného rozdělení dle vztahu:

$$X_i = F^{-1}_i \Phi_i \quad (11)$$

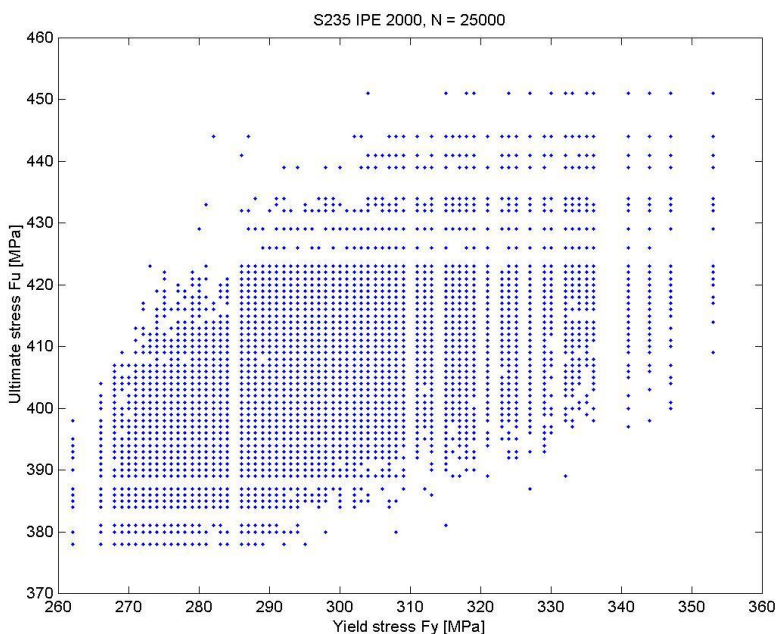
Distribuční funkce je zde získána prostým seřazením naměřených hodnot, přičemž každá hodnota má své pořadí dle velikosti. V souboru je 247 dvojic dat (f_y a f_u), a proto i inverzní distribuční funkce/řada má 247 hodnot. Náhodné realizace budou tedy z řady vybírány za pomoci rovnoměrného rozdělení generovaného předchozím postupem. Je patrné, že je třeba rovnoměrné rozdělení lehce zaokrouhlit tak, aby jeho výstupem byla celá čísla v rozsahu 1-247. Výsledné histogramy meze kluzu f_y a meze pevnosti f_u získané na základě 25 tis. realizací jsou zobrazeny na Obr. 7 resp. na Obr. 8. Vzájemná korelace uvedených veličin je vyobrazena na Obr. 9.



Obr. 7. Histogram meze kluzu f_y [MPa], 25 tis. simulací



Obr. 8. Histogram meze pevnosti f_u [MPa], 25 tis. simulací



Obr. 9. Simulovaná korelace mezi mezí kluzu f_y a mezí pevnosti f_u [MPa], 25 tis. simulací.

3.7. Ověření simulačního procesu

Pro ověření simulačního procesu je následně ověřena korelační matice $R_{p,k}$ náhodně generovaných korelovaných realizací a porovnána s původní maticí R_p . Pearsonův korelační koeficient je:

$$R_{p,k} = \begin{bmatrix} 1.0 & 0.649 \\ 0.649 & 1.0 \end{bmatrix} \quad (12)$$

Je vidět že rozdíl mezi korelačními součiniteli je v řádu 0,006, což je velmi povzbuzující vzhledem k počtu pouhých 25 tis. aplikovaných simulací.

Kvalita simulace histogramů je ověřena fraktilová korelací. Porovnáním shody ve frekvenci jednotlivých sloupců histogramů původních a simulovaných. Frekvence simulovaných je podělena počtem simulací a vynásobena 247 (počtem naměřených vzorků). Shoda je vidět již pouhým porovnáním obrázků (meze kluzu Obr. 1 a Obr. 7, resp. mez pevnosti Obr. 2 a Obr. 8). Pro mez kluzu i mez pevnosti je shodně pearsonův korelační koeficient roven 0.9995.

4. ZÁVĚR

Příspěvek, jež je založen na práci Phoona, at all [7], naznačuje možnosti generování korelovaných neparametrických rozdělení s využitím transformovaných korelovaných normálních rozdělení.

Ve studii je provedena analýza korelace meze kluzu f_y a meze pevnosti f_u . Data jsou získána z práce [9]. Na základě této korelace je vytvořena dvojice korelovaných normálních vektorů. Tyto jsou transformovány na dvě rovnoměrná korelovaná rozdělení, které slouží pro vytvoření obecného rozdělení (např. histogramu). Histogram je vytvořen standardním postupem inverzní transformací z rovnoměrného rozdělení dokumentovaným např. v [4].

Popis korelace je v této pilotní studii postaven na Pearsonovém korelačním koeficientu, který může být nahrazen Spearmanovým koeficientem pořadové korelace vhodnějším pro neparametrická rozdělení.

Uvedený postup umožňuje simulovat i větší množství korelovaných veličin a jeví se využitelný v rámci metody SBRA včetně přímé simulace technikou Monte Carlo.

Další vývoj by měl směřovat k hlubšímu prověření vhodnosti a efektivity transformačního procesu, testování možných aplikací (např. ocelové táhlo s otvorem) a po důkladném prověření k implementaci přístupu např. v rámci software Anthill.

Poděkování

Projekt byl realizován za finanční podpory ze státních prostředků prostřednictvím Grantové agentury České republiky. Registrační číslo projektu je GA ČR 103/07/0557.

Literatura

- [1] ANTHILL [on-line] dostupné na [www HTTP://WWW.SBRA-ANTHILL.COM/](http://www.sbra-anthill.com/).
- [2] VOŘECHOVSKÝ, M. and NOVÁK, D. (2002) Correlated random variables in probabilistic simulation. In Schiebl, Gebbeken, Keuser, and Zilch (eds), 4th International Ph.D. Symposium in Civil Engineering held in Munich, Germany, volume 2, pages 410-417. Millpress, Rotterdam, 2002.
- [3] MAREK, P., BROZZETTI J., GUŠTAR, M., TIKALSKY, P., EDITORS. Probabilistic Assessment of Structures using Monte Carlo Simulation. Basics, Exercises, Software. (Second extended edition)., Publisher: ITAM Academy of Sciences of Czech Republic, Prosecká 76, 190 00 Prague 9, Czech Republic, 2003 ISBN 80-86246-19-1.
- [4] MAREK, P., GUŠTAR, M., BATHON, L. Simulation-Based Reliability Assessment for Structural Engineers. Boca Taton, Florida, CRC Press, 1995, ISBN 0-8493-8286-6.
- [5] MATLAB [on-line] dostupné na [www HTTP://WWW.MATLAB.COM/](http://www.matlab.com/).
- [6] MENČÍK, J. (2003) Simulační posuzování spolehlivosti při korelovaných veličinách. in Sborník 4. Konference „Spolehlivost“, Ostrava, 23.-24.4. 2003, DT Ostrava, ISBN 80-02-01551-7, s. 151-156.
- [7] PHOON, K.,K. (2004) APPLICATION OF FRACTILE CORRELATIONS AND COPULAS TO NON-GAUSSIAN RANDOM VECTORS, IN CD-ROM PROCEEDINGS OF THE 2.ND INTERNATIONAL ASRANET COLLOQUIUM (5-7 JULY 2004), BARCELONA, SPAIN.
- [8] PHOON, K.,K., QUEK, S., T., HUANG, H., Simulation of non-Gaussian Processes using fractile correlation, in Probabilistic Engineering Mechanics, vol 19, p. 287-292, 2004.
- [9] ROZLÍVKA, L., FAJKUS, M. (2003) *Reálné pevnostní hodnoty konstrukčních ocelí a rozměrové úchytky válcovaných materiálů pro pravděpodobnostní posuzování spolehlivosti ocelových nosných prvků a konstrukcí metodou SBRA*;: Spolehlivost konstrukcí – Sborník referátů, Dům techniky Ostrava, Ostrava, 2003 ISBN 80-02-01551-7
- [10] VOŘECHOVSKÝ, M. (2007) *Stochastic computational mechanics of quasibrittle structures*. Habilitation thesis presented at Brno University of Technology, Brno, Czech Republic,
- [11] WEISSTEIN, E. W. "Correlation Coefficient." From MathWorld--A Wolfram Web Resource. [on-line] dostupné na [www HTTP://MATHWORLD.WOLFRAM.COM/CORRELATIONCOEFFICIENT.HTML](http://mathworld.wolfram.com/CorrelationCoefficient.html).
- [12] WEISSTEIN, E.W. "Erf." From MathWorld--A Wolfram Web Resource. [on-line] dostupné na [www HTTP://MATHWORLD.WOLFRAM.COM/ERF.HTML](http://mathworld.wolfram.com/Erf.html).
- [13] IMAN, R.C. AND CONOVER, W. J., A Distribution Free Approach to Inducing Rank Correlation Among Input Variables, *Journal Communications in Statistics*, 1982, vol.B11, pp. 311-334.

5. PŘÍLOHA: GENEROVÁNÍ KORELOVANÉHO NORMÁLNÍHO ROZDĚLENÍ V PROGRAMU ANTHILL [1]

Na podkladě korelační matice R (viz. část 3.2201 *Pearsonův korelační součinitel*) a postupu uvedeného v 3.4 *Generování korelovaných normálních rozdělení* jsou vygenerována dvě vzájemně korelovaná normální rozdělení, a to v programu Anthill. Je aplikováno 25 tis. simulací.

5.1. Vstupní parametry

Do výpočtu vstupují dvě vzájemně nezávislá useknutá normální rozdělení $N(0;1)$ W_1 a W_2 dále matice $Q = \text{chol}(R)$ (viz. (8)) jejíž řád odpovídá počtu proměnných.

5.2. Korelace normálních rozdělení

Korelace je mezi vzájemně nezávislé veličiny vnesena vyřešením soustavy rovnic reprezentované:

$$Y_N = W \times Q = \begin{bmatrix} w_1 & w_2 \end{bmatrix} \times \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} = \begin{bmatrix} w_1 \times q_{11} + w_2 \times q_{21} & w_1 \times q_{12} + w_2 \times q_{22} \end{bmatrix} \quad (13),$$

kdy Q je převzato z (8) a je:

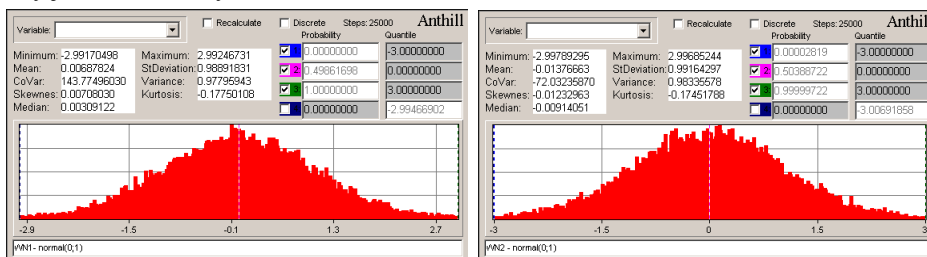
$$Q_{[k,k]} = \text{chol}(R) = \begin{bmatrix} 1.000 & 0.660 \\ 0.000 & 0.751 \end{bmatrix} \quad (8)$$

pak vychází že:

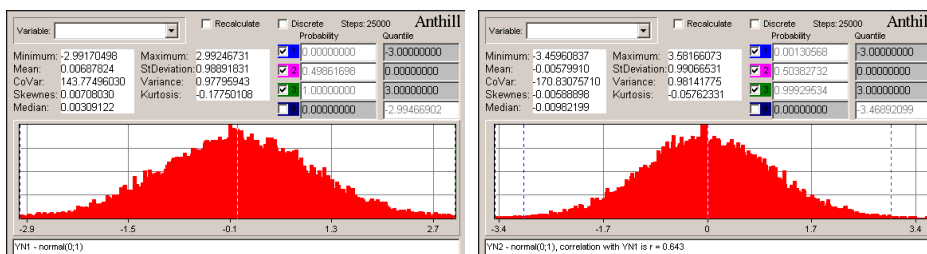
$$\begin{aligned} y_{N1} &= w_1 \times q_{11} + w_2 \times q_{21} = w_1 \times 1.0 + w_2 \times 0.0 = w_1 \\ y_{N2} &= w_1 \times q_{12} + w_2 \times q_{22} = w_1 \times 0.66 + w_2 \times 0.751 \end{aligned} \quad (14)$$

5.3. Výstupy z programu Anthill

Uvedená procedura 25 tis. opakována programem Anthill. Na následujících výstupech jsou ilustrovány jednotlivé kroky.

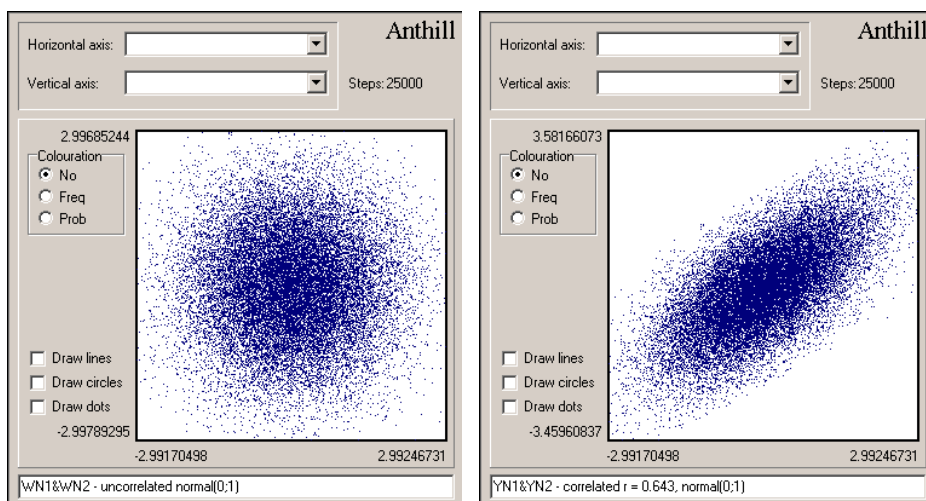


Obr. 10. Histogramy nekorelovaných normálních rozdělení w_1 a w_2 - $N(0;1)$



Obr. 11. Histogramy korelovaných normálních rozdělení y_{N1} a y_{N2} - $N(0;1)$, parametr korelace $r = 0.643$

Obr. 10 zobrazuje nekorelované normální rozdělení. Obr. 11 představuje předchozí normální rozdělení po transformaci (13) a Obr. 12 zobrazuje 2D diagram (mraveniště) vzájemného vztahu dvojic rozdělení.



Obr. 12. Mraveniště nekorelovaných normálních rozdělení w_1 , w_2 a korelovaných normálních rozdělení y_{N1} , y_{N2} pro 25 tis. simulací.

5.4. Souhrn a závěry

Zvolený postup, založený na obecně známých postupech (viz. např. v [7]) umožňuje přehledně generovat korelovaná normální rozdělení za pomoci simulační techniky Monte Carlo v rámci programu Anthill [1].

V příkladu je demonstrována korelace dvou proměnných, ale zdá se že aplikace na 3 a více proměnné je rovněž možná.

Generování bylo otestováno i pro 1 mil simulačních kroků. Proces trval na PC Celeron 300 MHz asi 6 min. Zdá se tedy, že proces není nikterak výpočtově náročný.

V Anthillu byla použita useknutá normální rozdělení a patrně proto je zřetelný jistý rozdíl mezi minimálními a maximálními hodnotami veličinami w_2 a y_{N2} . Zatímco w_2 je původní useknuté rozdělení, y_{N2} je korelováno z w_1 a w_2 . Tímto procesem došlo k rozmazání (zvětšení) minimální a maximální hodnoty.

Oponentura: Doc. Ing. Miroslav Vořechovský, Ph.D.

